The Asymptotical Error of Broadcast Gossip Averaging Algorithms

Fabio Fagnani **Paolo Frasca**

Dipartimento di Matematica

Politecnico di Torino, Italy

IFAC World Congress
Milan, September 1, 2011

Averaging in networks: Challenges

Distributed averaging is a building block to solve estimation problems in sensor and control networks.

Depending on the application, in distributed averaging we need to

- design efficient algorithms with little communication requirements
- analyze their performance
  - in terms of both speed and **accuracy**
  - as a function of the network topology and size (large networks)

Averaging in networks: Formal problem statement

Set-up:

- a set of nodes $\mathcal{V}$ of cardinality N
- a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (undirected in this talk)
- data at the nodes: $y_v \in \mathbb{R}$ for all $v \in \mathcal{V}$

Goal: estimate the average $y_{\text{ave}} = N^{-1} \sum_{v \in \mathcal{V}} y_v$.

Constraint:

- avoid synchronous node updates
- use instead directional asynchronous communication

$\longrightarrow$ use randomized broadcast communication

Averaging in networks: Formal problem statement

Set-up:

- a set of nodes $\mathcal{V}$ of cardinality N
- a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (undirected in this talk)
- data at the nodes: $y_v \in \mathbb{R}$ for all $v \in \mathcal{V}$

Goal: estimate the average $y_{\text{ave}} = N^{-1} \sum_{v \in \mathcal{V}} y_v$.

Constraint:

- avoid synchronous node updates
- use instead directional asynchronous communication

$\longrightarrow$ use randomized broadcast communication

Broadcasting Gossip Algorithm (BGA): definition

Broadcast Gossip Algorithm

for $v \in \mathcal{V}$ **do**
  $x_v(0) = y_v$
end for
for $t \in \mathbb{Z}_{\geq 0}$ **do**
  Sample node v from a uniform distribution over $\mathcal{V}$
  for $u \in \mathcal{V}$ **do**
    if $u \in \mathcal{N}_v$ **then**
      $x_u(t+1) = (1-q)x_u(t) + q\, x_v(t)$
    else
      $x_u(t+1) = x_u(t)$
    end if
  end for
end for

Mixing parameter:
$q \in (0, 1)$

A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione. Gossip algorithms for distributed signal processing. *Proceedings of the IEEE*, 98(11):1847–1864, 2010

Preliminary results & definitions

Proposition (Convergence)

If $\mathcal{G}$ is connected, then there exists a random variable x^* such that almost surely $\lim_{t\to+\infty} x(t) = x^*\mathbf{1}$.

Proposition (Martingale property)

Let $x_{\text{ave}}(t) = N^{-1}\sum_{v\in\mathcal{V}} x_v(t)$.
Then $\{x_{\text{ave}}(t)\}_t$ is a martingale (w.r.t. $x(t)$) and $\mathbb{E}[x^*] = x_{\text{ave}}(0)$.

However, x^* is not equal to $x_{\text{ave}}(0)$. The goal of this work is studying

$$\beta(t) = |x_{\text{ave}}(t) - x_{\text{ave}}(0)|^2,$$

and in particular the limit $\mathbb{E}[\beta(\infty)] := \lim_{t\to\infty} \mathbb{E}[\beta(t)]$

Preliminary results & definitions

Proposition (Convergence)

If $\mathcal{G}$ is connected, then there exists a random variable x^* such that almost surely $\lim_{t\to+\infty} x(t) = x^*\mathbf{1}$.

Proposition (Martingale property)

Let $x_{\text{ave}}(t) = N^{-1}\sum_{v\in\mathcal{V}} x_v(t)$.
Then $\{x_{\text{ave}}(t)\}_t$ *is a martingale (w.r.t.* $x(t)$*) and* $\mathbb{E}[x^*] = x_{\text{ave}}(0)$.

However, x^* is not equal to $x_{\text{ave}}(0)$. The goal of this work is studying

$$\beta(t) = |x_{\text{ave}}(t) - x_{\text{ave}}(0)|^2,$$

and in particular the limit $\mathbb{E}[\beta(\infty)] := \lim_{t\to\infty} \mathbb{E}[\beta(t)]$

Main result

Bounding the error introduced at each time step as

$$|x_{\text{ave}}(t+1) - x_{\text{ave}}(t)| \leq q\,\frac{d_{\max}}{N} L, \qquad \text{where} \quad L \geq \max_{u,v} |x_u(0) - x_v(0)|,$$

and exploiting the martingale property, we can prove

Theorem (Uniform Mean Square Error Bound)

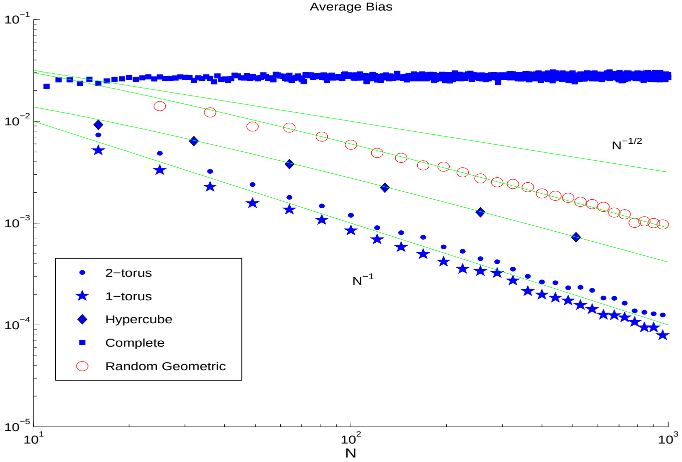
Let $\mathcal{G}$ be connected, λ_1 be its spectral gap and $d_{\max}$ be the maximum degree of its nodes. Then,

$$\mathbb{E}\left[\sup_{t\in\mathbb{N}} \beta(t)\right] \leq 8\,L^2 \frac{q}{1-q}\,\frac{d_{\max}^2}{N\lambda_1}$$

Think of large networks....

Simulations: $\beta(\infty)$ vs N

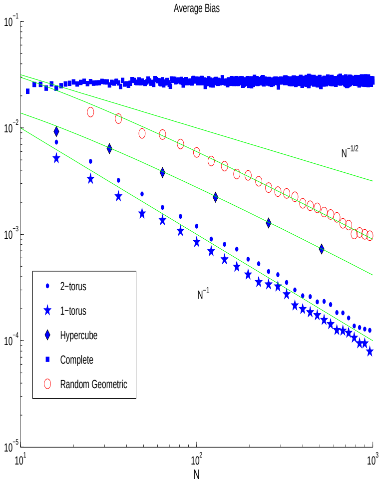
Varying size in example sequences



Solid lines are proportional to $N^{-1/2}$, $\log N/N$, $\log N/N$ and N^{-1}

Simulations: $\beta(\infty)$ vs N

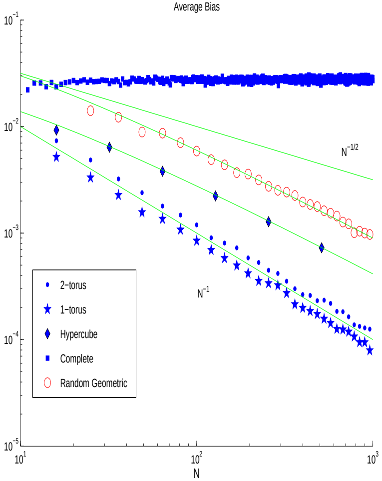
Varying size in example sequences



Solid lines are proportional to $N^{-1/2}$, $\log N/N$, $\log N/N$ and N^{-1}

Simulations: $\beta(\infty)$ vs N

Varying size in example sequences



graph	degree	$\beta(\infty)$
complete	$N-1$	constant
ring (1-torus)	2	$1/N$
2-torus	4	$\sim 1/N$
hypercube	$\log N$	$\sim \frac{\log N}{N}$
random geometric	$\sim \log(N)$	$\sim \frac{\log N}{N}$

Solid lines are proportional to $N^{-1/2}$, $\log N/N$, $\log N/N$ and N^{-1}

Summary and Further Research

In the BGA, a **larger network** gives a **more accurate** averaging!

Simulations suggest $\propto \dfrac{d_{\max}}{N}$ **but** we have proved $\propto \dfrac{d_{\max}^2}{N\lambda_1}$

Future research

- Close this gap!

  Related known result: for 1- and 2-dimensional tori, $\lim_{N\to\infty} \beta(\infty) = 0$

  F. Fagnani and P. Frasca. Broadcast gossip averaging: interference and unbiasedness in large Abelian Cayley networks. *IEEE Journal of Selected Topics in Signal Processing*, 5(4):866–875, 2011

- Extend this analysis to other randomized averaging algorithms

Further related reading

About the BGA and its performance:

F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE Journal on Selected Areas in Communications*, 26(4):634–649, 2008

T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione. Broadcast gossip algorithms for consensus. *IEEE Transactions on Signal Processing*, 57(7):2748–2761, 2009

T. C. Aysal, A. D. Sarwate, and A. G. Dimakis. Reaching consensus in wireless networks with probabilistic broadcast. In *Allerton Conf. on Communications, Control and Computing*, pages 732–739, Monticello, IL, September 2009

A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Transactions on Automatic Control*, 55(1):225–230, 2010